

“トモグラフィック表示”による 画像の局所的フーリエ解析

Local Spatial Fourier Analyses of Images
using "Tomographic Representation"

岡島 健治 藤原 正三
K. Okajima S. Fujiwara

日本電気(株) 基礎研究所
NEC Corp. Fundamental Research Labs.

Abstract A mathematical model of the primary visual cortex is proposed, and its functions are analyzed from the viewpoint of visual information processing. According to the model, local images are represented using a certain representation("tomographic representation") and Fourier analyzed. It is shown that 1) the proposed processing is tolerant to shifts in position of input images, and that 2) spatial frequency filtering operations can be easily performed in the model.

1. はじめに

まず始めに、『工学上の種々の問題に対して、生体の持つ優れた機能を模倣し、応用する』という手法の有効性について考えてみたい。

脳・中枢神経系は例えば、高度の並列処理機能、或いは学習機能等の多くの優れた機能を持っていると考えられている。そして、それは現在のディジタルコンピュータが苦手とする、例えばパタン処理、連想記憶等の面に於いて優れた能力を有しているというのが今日の一般的な認識であろう。従って、これらの面において、生体における情報処理のメカニズムを解明し模倣する事ができれば、全く新しい技術の展開が期待できるものと考えられる。

所で、周知の様に、1960年代にも全く同様な考え方に基づいてバイオニクス研究が展開されたのだが、それはパタン認識等の応用面において、必ずしも当初の期待を満足させる結果を生まなかった。これには、いくつかの理由が考えられるが、中でも次のような点が挙げられよう。

●実証的な脳研究のレベルがまだ不十分であった。それにも拘らず、安易に応用へと結びつけようとする考え方に

無理があった。

●モデルをハードウェア化するための適当なデバイスがなかった。

●デジタルコンピュータを用いての、他の方式による装置が、コスト的にも安く、ある程度の機能を実現させてしまった。

これに対して現状を考えるに、例えば現在のパタン認識装置、或いは画像理解技術等と人間の持っているパタン認識能力との間には、大きいギャップがある事は否めず、なんらかの突破口が必要とされている状況である。次に特に視覚情報処理系に関して言えば、Hubel, Wiesel⁽¹⁾の先駆的な研究を端緒として多くの実証的な研究がなされ脳に関する知識の蓄積がハイスピードで進んでいる。さらにデバイスに目を向ければ、集積化技術の進展により、その加工精度は、既に神経細胞の大きさの程度に達してきている。

以上の様な点を考慮するならば、現時点において再び、「生体のメカニズムに学ぶ」という手法に挑戦する事は十分に意義がある事であろうと考えられる。

また、ここで視点を変えて、脳の解

明という視点から考えてみよう。従来それは、神経生理学、心理学等の実証的な手法により主として進められて来た。これらの実証的なアプローチが脳の解明の為には第一義的に重要である事は言うまでもない。しかしながら、脳という器官は、情報処理（それも、かなり複雑な）を仕事とする器官であるのだから、その解明を進めるにあたっては情報処理という視点に立った工学サイドからの寄与が不可欠である事も容易に推測できる。

さて、ここでは以上の様な考え方に基づく立場から、視覚系を取り上げてその初期の部分（視覚一次野）に於いてなされる視覚情報処理過程を議論したい。

ここで、注意を要する点は、視覚系それ自体もまたいくつかのサブシステムから成り立つシステムであるという点である。従って、その一部に関する知見に基づいた応用に対して過大な期待をする事は誤りであり、むしろ、問題を注意深く分解して行き、根気強く取り組む努力が必要となろう。この点に関して少し具体的に説明を加えたい

図1はサルの大脳に於ける、視覚系の概略を示している。網膜から入力した視覚情報は、中継核（外側膝状体）を介して、後頭部にある視覚一次野に投写される。ここで処理を受けた視覚情報は、その後大きく分けて二つの経路にそって大脳皮質上で処理を受けると考えられている⁽²⁾。一つは、視覚一次野から視覚領前野を経て頭頂連合野に向かう、上縦束と呼ばれている経路である。この経路は空間視に関する情報処理を担っていると考えられている。例えば、対象物を構成する要素の間の空間的相互関係、あるいは、対象物と自分自身との間の空間的相互関係等の認識などの処理、さらには協調的な眼球運動による対象全体の認識処理

などが行なわれていると考えられている^(2,3)

第二の経路は、視覚一次野から視覚領前野を経て側頭連合野へ向かう下縦束と呼ばれる経路である。この経路は視対象そのものの性質に関する情報処理を担っていると考えられている（形態視、及びその記憶）。視覚一次野の神経細胞が体局在性を示し、各々網膜上の特定の局在領域を担当しているのに対して、側頭連合野では網膜上の広い範囲にわたる領域の情報が一つの細胞に収められている事が知られている⁽²⁾。即ち、視覚一次野等で局所的に処理された情報は側頭連合野に於いて広い範囲にわたって統合され知覚されていると考えられる。

以上のような、例えば、空間視が関係する対象の認識、或いは広い領域にわたる情報の統合等の機能は、さらに

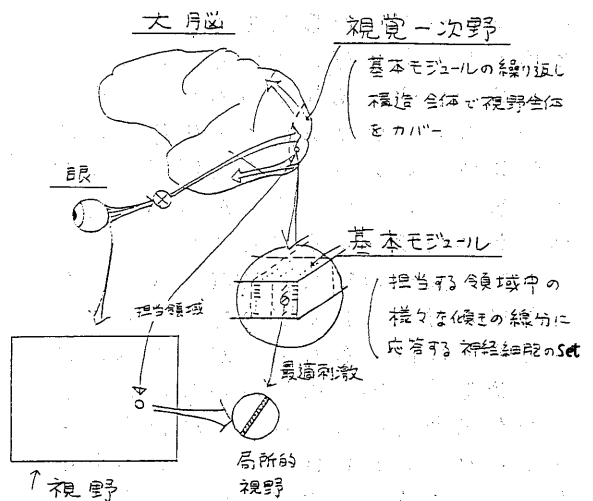


図1. 視覚経路と視覚一次野

高次の機能になり、視覚一次野の機能とは別に考える必要があるのである。

以下、2に於いて視覚一次野に関する生理学的、及び心理学的知見をいくつか述べ、3でその数学的モデルを提案する。そして、4に於いて、このモデルに関して、その動作、機能に関して、工学的な視点からの検討を加えたい。最後に5で結論をまとめる。

2. 生理学、心理学的知見

Hubel, Wiesel らによる先駆的な研究、及びそれに続く一連の電気生理学的研究によって、今日では視覚一次野に関して多くの知見が得られている⁽¹⁾。その主要な点をまとめると以下のようになる(図1)。

●1) 視覚一次野はハイパーコラムと呼ばれる基本的モジュール(大きさ1-2mm²程度)の繰り返し構造により構成されている。

●2) 各基本モジュールは各々網膜上の特定の小領域に投影された像の処理を担当している。各基本モジュールの担当領域は、モジュールごとに異なっており視覚一次野中の全て基本モジュールによって視野全体がカバーされている。

●3) 視覚一次野中の多くの神経細胞は、視野中に呈示された特定の傾きの線分に対して最大の応答を示す。さらに、これらの神経細胞のうち単純型細胞は特定の位置に呈示された線分に対してのみ応答を示すのに対して、複雑型細胞(及び超複雑型細胞)は、線分の傾きさえ正しければ、それが呈示される受容野中の位置によらずに応答を示す。

●4) これらの細胞は各基本モジュール中に整然と配置されており、各基本モジュールには、あらゆる傾きの線分をカバーする様な細胞群の完全なセットが含まれている。

問題は、このような装置が脳に於ける視覚情報の解析プロセスに於いて、どのような機能を果たしているのかという点にある。従来、これらの神経細胞は網膜上に投影された像中から線分という特徴を抽出している特徴抽出細胞であるとする見方が一般的であった。確かに線分は自然画像を表現する上で重要な特徴である事には間違いない。しかしながら、線分だけでは、自然画像を記述するための特徴としては十分ではないだろうから、この様な解釈には問題が残る。

他方、Campbell, Robsonらは^(4,5)主に心理物理的な実験結果に基づいて、人間の視覚系には、像の空間周波数分析を行なっている神経機構があるはずだとする仮設(マルチチャンネル仮設)を提案している。もしも彼らの仮設が正しいとすると、

●何の為に視覚系は空間周波数分析を行なっているのか

●脳のどの部分がそれを行なっているのか

という点が問題となって来る。

3. 視覚一次野の数学的モデル

ここで提案するモデルは基本的には二つの特徴からなっている。第一に、このモデルでは、各基本モジュール中の単純型細胞は、単に画像中の線分を特徴として抽出しているのではなく、むしろ全体として、その基本モジュールが担当する局所像を特定の表示(ここではそれをトモグラフィック表示と呼ぶ事にする)によって表現しているのだと仮定する。第二に、このモデルでは各基本モジュールは、その担当する局所像に対して空間フーリエ解析を行なっているものと仮定する。従って視覚一次野全体としては入力画像の局所的空間フーリエ解析を行なっていることになる。以下にこれらの仮定に関して、やや詳細に説明を加える。

3-1. トモグラフィック表示

図2にCT (Computerized Tomography) の原理図を示した。例えば、X線CT等では、細く絞ったビームをスキャンさせながら対象に照射し、検出器によって透過量を測定して、ビームの減衰量を求める。このようなスキャンを角度を変えながら実行すると、得られたデータから元の物体の密度分布を再現することができる。

今、画像の明暗度を $a(r, \phi)$ としてそのトモグラフィックな表示 $d(t, \theta)$ を次式のように定義する。

$$d(t, \theta) = \iint a(r, \phi) \delta(t - r \cos(\theta - \phi)) |r| dr d\phi, \quad \delta(x): \text{デルタ関数}$$

CTの場合には、 $a(r, \phi)$ は物体の密度分布に、 $d(t, \theta)$ はX線の減衰量に対応する。

画像の明暗度、 $a(r, \phi)$ は、そのトモグラフィックな表示、 $d(t, \theta)$ から次式によって再現することができる^(6,7)。

$$a(r, \phi) = \int_0^\pi g(r \cos(\phi - \theta), \theta) d\theta$$

$$; g(t, \theta) = \int d(\tau, \theta) h(t - \tau) d\tau$$

$$, h(t) = \int_{-Kc}^{Kc} k | \exp(ikt) dt$$

従って、もしも、各単純型細胞の活動度が、この $d(t, \theta)$ を表わしているものと考えたと、基本モジュール中の全ての単純型細胞の興奮のパターンから、その基本モジュールが担当する局所画像が再現できるはずである。その意味でこの表示は完全である。

勿論、トモグラフィック表示の基底は様々な傾きの線分からなるのだからこれは元来線分を扱うのに適した表示である。しかしながら重要な点は、それが完全な表示であるために、線分以外の特徴をも処理する事ができると

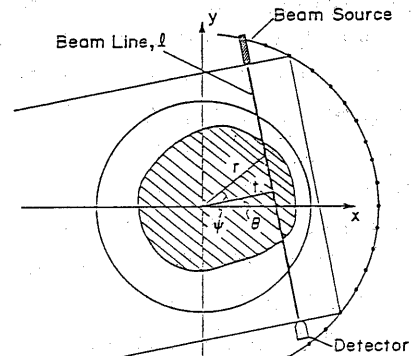
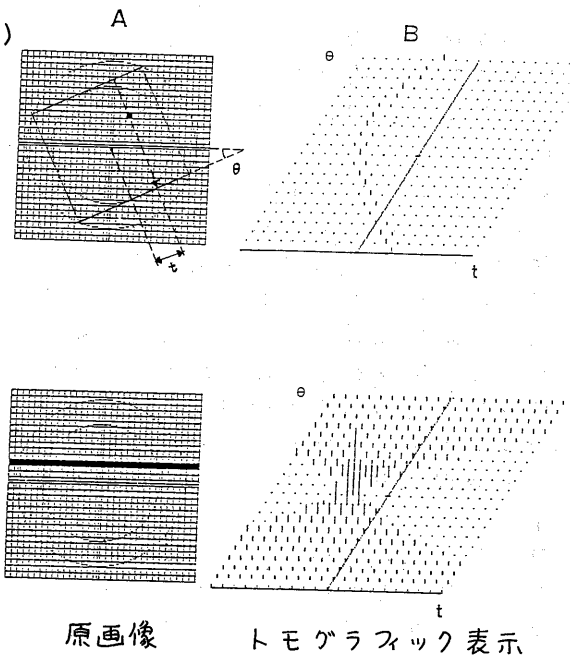


図2. コンピュータトモグラフィの原理図

いう点である。この点が、従来のように、視覚一次野の細胞を画像中の線分という特徴を抽出している特徴抽出細胞であるとする見方との大きな相違点である。

図3にトモグラフィック表示の例を示した。



原画像 トモグラフィック表示

図3. トモグラフィック表示の例

3-2. 局所的空間周波数解析

画像をトモグラフィック表示で表現する事の一つのメリットは、画像のフーリエ変換、或いは相関関数の計算等の2次元の処理が1次元の処理によって実現できる点にある。

今画像、 $a(r, \phi)$ をトモグラフィックな表示、 $d(t, \theta)$ に変換しておいて変数 t に関してフーリエ変換を施してみる。

$$\begin{aligned} D(k, \theta) &= \int d(t, \theta) \exp(-ikt) dt \\ &= \iiint a(r, \phi) \delta(t - r \cos(\theta - \phi)) |r| dr d\phi e^{-ikt} dt \\ &= \iint a(r, \phi) \exp(-ik \cdot r \cos(\theta - \phi)) |r| dr d\phi \\ &= \iint a(r) \exp(-ik \cdot r) dr = A(k) \\ &, k = k(\cos \theta, \sin \theta) \end{aligned}$$

即ち、トモグラフィックな表示、 $d(t, \theta)$ の添え字 t に関するフーリエ変換、 $D(k, \theta)$ は、元の画像 $a(r)$ の添え字 r に関するフーリエ変換、 $A(k)$ を極座標表示したものに等しい。

このように、画像をトモグラフィックな表示で表現しておく、と一次元のフーリエ変換によって、元の像の二次元のフーリエ変換を求めることができるのである。

トモグラフィックな表示に於いて、フーリエ変換を計算する為には、同じ傾き選択性を持ち、位置選択性が異なる、基本モジュール中の全ての単純型細胞の活動度を一つの細胞に集める必要がある ($d(t, \theta)$; $\theta = \text{const.}$)。ここではある種の複雑型細胞がその役割を果たしていると仮定する。この仮定はGlezerらによる電気生理学的実験結果と整合している⁽⁸⁾。

以上述べたように、このモデルではCampbell, Robsonらの提案した周波数チャンネルは、ある種の複雑型細胞によって実現されていると仮定されている。ただし、この場合の周波数チャ

ネルは、局所的空間周波数に対するチャンネルとなる。

(変換の例) 図4に、画像(図9Bと同じもの)を小領域ごとに(トモグラフィック表示にして)フーリエ変換を行なって、空間周波数スペクトルを求めた例を示した。原画像は、 200×256 画素の2値画像である ($a(x, y)$)。これを 13×10 コの小領域に分割して、($a(x, y; X, Y)$) トモグラフィック表示にしてから ($d(t, \theta; X, Y)$)、添え字 t に関するフーリエ変換を求める ($A(k, \theta; X, Y)$)。ここではその振幅成分(空間周波数スペクトル ($P(k, \theta; X, Y)$))のみが示されている。また、領域の境界での不連続性を消すためにフーリエ変換の際には、分散 $\sigma = 10$ (画素) のガウシアン窓を用いた。

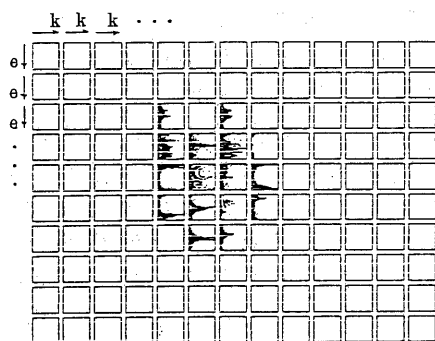


図4. モデルによる処理の例

4. モデルの機能の検討

4-1. 位置ずれに対する安定性

良く知られているように、画像の空間周波数パワースペクトルは入力画像の位置には依存しないから、このモデルでの処理の出力(振幅成分)は小領域のサイズ程度の位置ずれに対しては安定である(図5)

原画像(図9Bに示したもの)を $\Pi (= a(x, y))$ として、それを処理したものを、 $\Theta (= P(k, \theta; X, Y))$ で表

わす。

ここで、原画像、 Π を $(\delta X, \delta Y)$ だけ変位させたものを Π' とし、それを処理した結果を、 Φ' とする。

図5には、 Π と Π' 、 Φ と Φ' の類似度

$$\Pi \cdot \Pi' / |\Pi| |\Pi'| \quad (\text{図5 A}) \quad \text{及び}$$

$$\Phi \cdot \Phi' / |\Phi| |\Phi'| \quad (\text{図5 B})$$

が変位の関数として示されている。

原画像では、画像間に数画素程度の変位があると、類似度は急激に減少していく。それに対して、原画像間に（小領域の大きさ程度の）位置ずれがあっても、処理後のパターンでは、類似度はある程度大きく、位置ずれに対して安定になっていることがわかる。

（分解能）このような、位置ずれにたいする安定化は、適当な重み関数によって、画像をぼかす事によっても実現できる。適当な重み関数（ガウシアン等）によって移動平均をとると、重み関数の巾程度の範囲で、パターンがぼける為、（重み関数の巾、 σ 程度の範囲の）位置ずれに対して、パターンは安定になる。しかし、この場合には、移動平均をとる過程で、平滑化が行なわれて、パタンの空間周波数の $(\sigma_K = 1/\sigma \text{ 以上の})$ 高周波成分が、捨てられてしまうために分解能が落ちてしまう。重み関数の巾を広げるとパターンは位置ずれに対して、より安定になるが分解能は悪くなる。逆に重み関数の巾を狭めると分解能は良くなるが、位置ずれに対する安定度は悪くなる。このように、移動平均によりパターンをぼかす方法では、許容される位置ずれの大きさと分解能とは、トレードオフの関係にある。それに対して、小領域に分割して空間周波数分析を行なう処理では、パタンの持つ高周波の情報も処理の過程に於いて保存されるので、位置ずれに対する安定化を、分解能を保ったまま実現できるというメリット

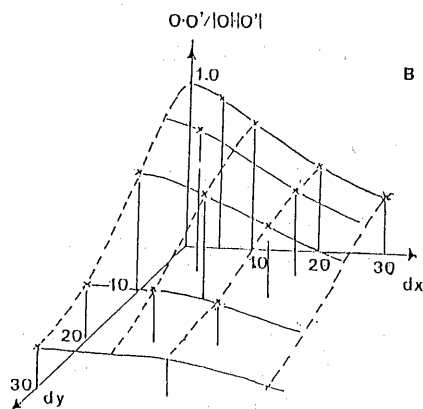
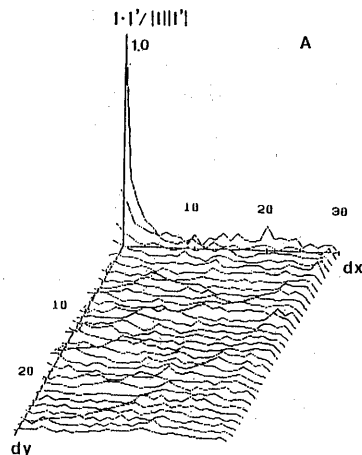


図5. 変位に対する安定性

がある。

図6は、1本の直線 $(0 < y < 2)$ と2本の直線 $(0 < y < 1, 2 < y < 3)$ とを示している。もしも、位置ずれに対して安定化させるために、これらのパターンを $\sigma \sim 1$ 以上のガウシアンでぼかしてしまうと、二つのパターン間の重なり合いはほぼ1となってしまう。これらの二つのパターンは、もはや区別できない。

それに対して、このモデルによる処理では、各小領域で処理されたパターン

の高周波成分によって、これらの二つを区別し得ることが分かる（図7）。このように、このモデルによる処理では、位置ずれに対する安定化を分解能を落とす事なく実現できる。

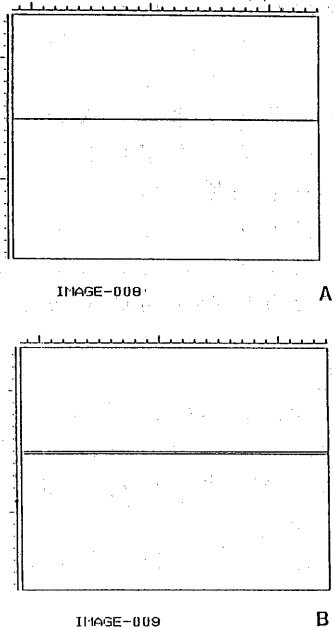


図6. 1本の直線と2本の直線

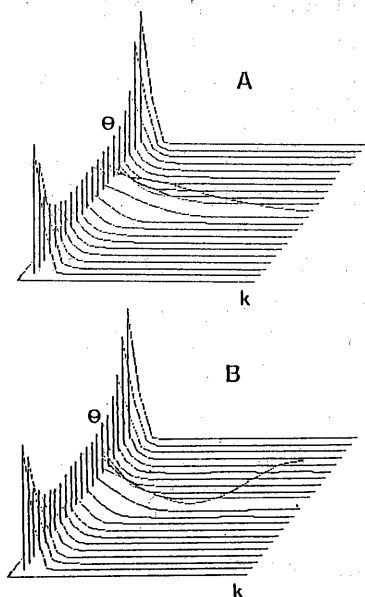


図7. 基本モジュールによる
図6の処理結果

4-2. 周波数空間でのフィルタリング操作

図8は、Harmonによるもので、画像に量子化処理を施したものを示している⁽⁹⁾。これらの図を普通に見ていると量子化のために生じる正方形のエッジまでもが見えてしまい、そのために、像の認識が妨げられてしまう。ところが、例えば、目を細めてわざと分解能を落とし、像の高周波成分を捨ててやって眺めると、ケネディー（左）やチャーチルの顔（右）等が認識できるようになる。

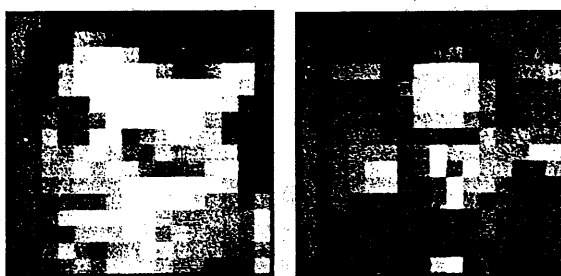


図8. 量子化された画像（9）

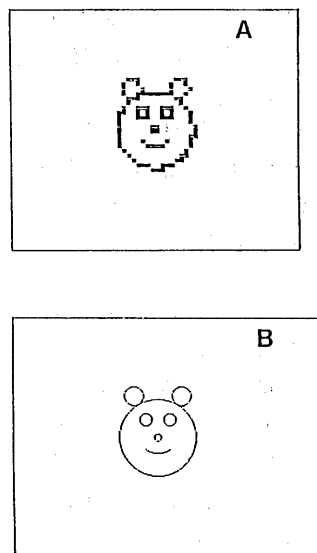


図9. 画像（B）とその量子化された画像（A）

ところで、3で説明したようなモデルでは、このような周波数空間でのフィルタリング操作が容易に実現できる（例えば高周波成分を担当する神経細胞に抑制をかけることによって）。

図9Bのような図形と、それを量子化した図形（図9A、量子化の単位＝4×4画素）とを用いて、それらを3で提案したモデルに従って処理し、処理後のパターン同志の類似度を、周波数空間で適当な重みをつけて計算してみた（注1）。

Case-1. . . . 0. 5 2 3

Case-2. . . . 0. 9 8 7

（高周波成分
を抑制した場合）

このように、3で提案したモデルでは、注目するサイズに応じて適当な周波数を選択しての処理を行なう事ができ、例えば、図8に示したような二つの像を、同一（Case-2、高周波成分を抑制した場合）とも、異なる（Case-1）とも処理し得ることが分かる。

5. 議論及びまとめ

3で提案した視覚一次野のモデルをまとめると以下のとおりである。

- 入力画像はいくつもの小領域に分割され、各々、各ハイパーコラムによって処理を受ける。
- 各ハイパーコラムの担当領域に投影された局所画像はそのハイパーコラム中の全ての単純型細胞を用いてトモグ

ラフィック表示によって、表現される。

- 各ハイパーコラムはこうしてトモグラフィック表示された局所画像に対しフーリエ解析を行う。従って、全画像は、視覚一次野全体によって、局所的空間周波数解析を受ける。

このようなモデルによる処理は次の様な特長を持っている。

- 分解能を落とす事なく入力パタンの位置ずれに対する安定化を実現できる。
- 空間周波数領域でのフィルタリング操作が容易であり従って、入力パターンに特徴的なサイズに応じた周波数を選択しての認識処理が可能である。

参考文献

- 1) Hubel, D.H., Wiesel, T.N.,
Proc. R. Soc. Lond. B198,1(1977)
- 2) Mishkin, M.: 生体の科学, 30(6)
437(1979)
- 3) Luria, A.P. " 神経心理学の基礎"
医学書院
- 4) Campbell, F.W. and Robson, J.G.
J. Physiol., 197, 51(1968)
- 5) Blakemore, C., Campbell, F.W.
J. Physiol., 203, 237(1969)
- 6) Radon, J., Ber. Sachs. Akad. Wiss.
69, 262(1917)
- 7) Hounsfield, G.N., British J.
Radiography, 121, 189(1976)
- 8) Glezer, V.D. et al.
Vision Res., 13, 1875(1973)
- 9) Harmon, L.D., Scientific
American 232 April 33(1975)

（注1）3で説明したように、トモグラフィックな表示、 $d(t, \theta)$ の添え字、 t に関するフーリエ変換、 $D(k, \theta)$ は、像 $a(r)$ のフーリエ変換、 $A(k)$ を極座標表示したものに等しい。従って、処理結果 $P(k, \theta; X, Y)$ と $P'(k, \theta; X, Y)$ とで、通常の内積 $\sum_{X,Y} \iint PP' dkd\theta$ を取ると、高周波成分が過小に評価されてしまう（極座標表示では、面積要素 dS は、 $dS = k dkd\theta$ である）。これが、Case-2である。重み関数として、 $M(k) = k \exp(-0.125k)$ を用いて $\sum_{X,Y} \iint PP' M^2 dkd\theta$ で、高周波成分の重みを増して、内積を計算したのがCase-1である。